AD 44789

AD No. 44 789
ASTIA FILE COPY

THE MOMENTS OF THE SAMPLE MEDIAN

by

J. T. Chu and Harold Hotelling

Special report to the Office of Naval Research of work at Chapel Hill under Contract NR 042 031, Project N7-onr-284(02), for research in Probability and Statistics.

Institute of Statistics
Mimeograph Series No. 116
August, 1954

THE MOMENTS OF THE SAMPLE MEDIAN[1,2]

by

John T. Chu and Harold Hotelling
Institute of Statistics, University of North Carolina

1. Summary. It is shown that under certain regularity conditions, the moments about its mean of the sample median tend, as the sample size increases indefinitely, to the corresponding ones of the asymptotic distribution (which is normal). A method of approximation, using the inverse function of the cumulative distribution function, is obtained for the moments of the sample median of a certain type of parent distribution. An advantage of this method is that the error can be made as small as is required. Applications to normal, Laplace, and Cauchy distributions are discussed. Upper and lower bounds are obtained, by a different method, for the variance of the sample median of normal and Laplace parent distributions. They are simple in form, and of practical use if the sample size is not too small.

2. Introduction. Let a population be given with cdf (cumulative distribution function) $F(x)$ and pdf (probability density function) $f(x)$, and median ξ which we assume to exist uniquely. Let $\tilde{x}$ denote the sample median of a sample of size $2n + 1$. Then the pdf $g(x)$ of $\tilde{x}$ and the pdf $h(x)$ of the asymptotic distribution of $\tilde{x}$ are respectively

$$g(x) = C_n [F(x)]^n [1 - F(x)]^n f(x), \tag{1}$$

where $C_n = (2n + 1)!/(n!\, n!)$, and

---

1. This paper was presented at the summer meeting of the Institute of Mathematical Statistics at Montreal, P.Q., Canada, on September 11, 1954.

2. Sponsored by the Office of Naval Research under Contract NR 042 031 for research in Probability and Statistics.

$$(2) \qquad h(x) = (2\pi)^{-1/2} e^{-(x-\xi)^2/(2\bar{\mu}_2)},$$

where $\bar{\mu}_2 = \{4[f(\xi)]^2(2n+1)\}^{-1}$.

A question that follows naturally is: Can the moments of the asymptotic distribution of $\tilde{x}$ be used as approximations to the corresponding moments of $\tilde{x}$, and if not, how to find better approximations? When the parent distribution is normal, this question has been answered by various authors, e.g., T. Hojo [6], K. Pearson [8, 9] and more recently, J. H. Cadwell [3]. It has been stated, e.g., in [3], that experiments showed that the distribution of $\tilde{x}$ tends rapidly to normality, but the variance of $\tilde{x}$ (as of quantiles in general) tends only slowly to the variance of the asymptotic distribution. For this reason of slow convergence, approximations were derived for the variance of $\tilde{x}$ when the sample size is small. While different methods were used by different authors, their results agree fairly well with each other. In fact, the problem should be considered as completely solved but for the unknown error committed in using such approximation.

One of us [4] recently proved that the distribution of $\tilde{x}$, for a normal parent distribution, does tend to normality rather "rapidly". In § 6 we shall confirm another experimental result that the variance of $\tilde{x}$ tends "slowly" to the variance of the asymptotic distribution (actually not "very slowly"). Upper and lower bounds are obtained (§ 6, Theorem 4) for the variance of $\tilde{x}$. A slightly better lower bound is obtained, by a different method, in § 5, formula (49), or § 6, formula (74). It seems that even for sample sizes around 10 to 20, the asymptotic variance is not a bad approximation to the true variance of $\tilde{x}$. It becomes a very good approximation if the sample size is large. However, for large samples, an even better approximation is obtained in § 5, formula (56).

Before further discussion, the following notations will be introduced. If $f(x)$ and $g(x)$ are functions of x, then $E_f(g)$ denotes the expectation of $g(x)$ with

respect to $f(x)$, i.e., $\int_{-\infty}^{\infty} g(x)\, f(x)\, dx$. We use, where f, g, and h are given by (1) and (2),

$$(3) \qquad \tilde{\mu}_1 = E_g(x) \quad , \qquad \bar{\mu}_1 = E_h(x) = \xi \ ,$$

$$\mu_1' = E_g(x - \bar{\mu}_1) \ , \qquad \mu_1 = E_f(x) \quad ,$$

and for any integer $k \geq 2$,

$$(4) \qquad \tilde{\mu}_k = E_g(x - \tilde{\mu}_1)^k \ , \qquad \bar{\mu}_k = E_h(x - \bar{\mu}_1)^k \ ,$$

$$\mu_k' = E_g(x - \bar{\mu}_1)^k \ , \qquad \mu_k = E_f(x - \mu_1)^k \ .$$

It should be pointed out that, although the pdf $g(x)$ of $\tilde{x}$ tends to $h(x)$, the moments $\tilde{\mu}_k$ of $\tilde{x}$ in general do not necessarily tend to $\bar{\mu}_k$. In fact $\tilde{\mu}_k$ may never exist [2]. Nevertheless, if the parent pdf satisfies certain conditions, then it can be shown that $\tilde{\mu}_k$ tends to $\bar{\mu}_k$ as sample size tends to infinity (§ 3, Theorem 1). Therefore under such circumstance, it is justifiable, at least for large samples, to use $\bar{\mu}_k$ as an approximation to $\tilde{\mu}_k$.

If the parent distribution satisfies certain conditions, a general method is obtained in § 4 for computing $\tilde{\mu}_k$, $k = 1, 2, \ldots$. The method is based on the Taylor expansion of $x(F)$, the inverse function of $F(x)$. For example, if $x(F) - \xi = \sum_{m=1}^{\infty} a_m (F - \frac{1}{2})^m$ converges for $0 < F < 1$, $a_m = O(2^m m^k)$ where $k \gtrless 0$ and $f(x)$ is symmetric with respect to $x = \xi$, then when $n > 2k + 3$,

(5) $$\tilde{\mu}_2 \sim \int_0^1 S_m^2 C_n F^n(1 - F)^n \, dF ,$$

where C_n is given by (1) and $S_m = \sum_{r=1}^{m} a_r(F - \frac{1}{2})^r$ (§ 4, Theorem 3). Error in such approximation can be computed, and it tends to 0 as m tends to oo. If the parent pdf is not symmetric, similar approximations can be obtained (§ 4, formula 26). Applications are given to the variances of the sample medians of Laplace and Cauchy parent distributions (§ 4, Examples 1 and 2).

Finally upper and lower bounds are derived in § 7 for the variance of $\tilde{x}$ of a Laplace parent distribution. It then can be seen that for estimating the mean of a Laplace distribution, the sample median is a "better" estimate than the sample mean, not only for large samples, but for small samples as well.

3. Large Sample Moments.

Lemma 1. If $0 \leq c \leq \frac{1}{2}$, then for m, n = 1, 2, ...,

(6) $$\int_{\frac{1}{2}-c}^{\frac{1}{2}+c} \left| u - \frac{1}{2} \right|^m u^n(1 - u)^n \, du = \left(\frac{1}{2} \right)^{m+2n+1} \int_0^{4c^2} t^{(m-1)/2} (1 - t)^n \, dt .$$

In particular, if $c = \frac{1}{2}$, and $C_n = (2n + 1)!/n!\, n!$, we have for fixed m,

(7) $$\int_0^1 C_n \left| u - \frac{1}{2} \right|^m u^n(1 - u)^n \, du = O(n^{-m/2}) ,$$

$$(8)\quad \int_0^1 C_n(u-\tfrac{1}{2})^{2m}u^n(1-u)^n du = (\tfrac{1}{2})^{2m}\frac{1\cdot 3\cdots(2m-1)}{(2n+3)(2n+5)\cdots(2n+2m+1)},$$

$$(9)\quad \int_0^1 C_n\left|u-\tfrac{1}{2}\right|^{2m-1}u^n(1-u)^n du = (\tfrac{1}{2})^m\frac{1\cdot 3\cdots(2n+2m-1)}{2\cdot 4\cdots(2n+2m)}\cdot\frac{(m-1)!}{(2n+3)(2n+5)\cdots(2n+2m-1)}.$$

These formulae are easily proved using transformations $v = \pm(u - \frac{1}{2})$, etc.

Theorem 1. Let a population be given with cdf $F(x)$ and pdf $f(x)$. Suppose that the median ξ of the given population exists uniquely and $f(\xi) \neq 0$, and $f'(x)$ exists and is bounded in some neighborhood of $x = \xi$. If $\tilde{x}$ is the sample median of a sample of size $2n + 1$, and $\tilde{\mu}_k$ and $\bar{\mu}_k$, as defined by (4), are respectively the k^{th} moment of $\tilde{x}$ about its mean and the corresponding one of its asymptotic distribution, then

$$(10)\qquad \lim_{n\to\infty} \tilde{\mu}_{2k-1} = \bar{\mu}_{2k-1},$$

$$(11)\qquad \lim_{n\to\infty} \tilde{\mu}_{2k}/\bar{\mu}_{2k} = 1, \qquad k = 1, 2, \ldots,$$

provided that in each case, $\tilde{\mu}_{2k-1}$ and $\tilde{\mu}_{2k}$ are finite for at least one n.

Proof. We will prove (11) as an illustration of the method we use. (10) can be shown in the same way. Obviously

$$(12)\qquad \tilde{\mu}_{2k} = \mu'_{2k} + \sum_{j=0}^{2k-1}\binom{2k}{j}(-1)^{2k-j}\mu_1'^{\,2k-j}\mu'_j,$$

where $\binom{2k}{j} = (2k)!/(j!(2k-j)!)$. We say that if $\tilde{\mu}_{2k}$ is finite for a certain

$n = n_0$, then $\tilde{\mu}_{2k}$ is finite for all $n \geq n_0$, and

(13) $\mu'_{2m-1} = O(n^{-m})$,

(14) $$\mu'_{2m} = \left(\frac{a_1}{2}\right)^{2m} \frac{1 \cdot 3 \cdots (2m - 1)}{(2n + 3)(2n + 5) \ldots (2n + 2m - 1)} + O(n^{-m-1/2}),$$

for $m = 1, 2, \ldots, k$,

where $a_1 = 1/f(\xi)$. On combining (12), (13), and (14), it follows that

(15) $$\tilde{\mu}_{2k} = \left(\frac{a_1}{2}\right)^{2k} \frac{1 \cdot 3 \cdots (2k - 1)}{(2n + 3)(2n + 5) \ldots (2n + 2k + 1)} + O(n^{-k-1/2}) .$$

Since $\bar{\mu}_{2k} = 1 \cdot 3 \cdots (2k - 1)\, \bar{\mu}_2^k$, where $\bar{\mu}_2$ is defined by (2), we have (11).

To complete the proof, it remains to establish (13) and (14). Now for example,

(16) $$\mu'_{2m} = \int_{-\infty}^{\infty} (x - \xi)^{2m} c_n [F(x)]^n [1 - F(x)]^n f(x)\, dx$$

$$= \int_{-\infty}^{a} + \int_{a}^{b} + \int_{b}^{\infty} = I_1 + I_2 + I_3 , \quad \text{say},$$

where $a < \xi$ and $b > \xi$ will be chosen later. For $0 \leq F \leq 1$, the function $F(1 - F)$ is non-negative, reaches its maximum $\frac{1}{4}$ at $F = \frac{1}{2}$, and is increasing for $0 \leq F \leq \frac{1}{2}$ and decreasing for $\frac{1}{2} \leq F \leq 1$. Let

(17) $$r = \max \{4F(a)[1 - F(a)],\ 4F(b)[1 - F(b)]\} ,$$

then $0 < r < 1$. Since $C_n = O(2^{2n} n^{1/2})$, it follows that

$$I_1 + I_3 = O(n^{1/2} r^n) \ . \tag{18}$$

On the other hand, if a and b are so chosen that, e.g., $F(b) - \frac{1}{2} = \frac{1}{2} - F(a) = c$ is small, then for $\frac{1}{2} - c \leq F \leq \frac{1}{2} + c$, $x(F)$, the inverse function of $F(x)$, is uniquely defined and may be expanded, by Taylor's method, into

$$x(F) - \xi = a_1(F - \tfrac{1}{2}) + R_2(F - \tfrac{1}{2})^2 \ , \tag{19}$$

where $a_1 = 1/f(\xi)$ and R_2 is the remainder. Substituting (19) for $x - \xi$ in I_2 of (16), it can be shown, using Lemma 1, that I_2 is equal to the RHS (right hand side) of (14). Combining this fact with (16) and (18), we obtain (14).

Regarding the above theorem, we make

Remark 1. A sufficient condition for $\tilde{\mu}_k$ being finite for some $n = n_0$ (hence all $n \geq n_0$) is that μ_k be finite. This condition, however, is not necessary. For example, the variance of the sample median of a Cauchy parent distribution is finite if the sample size $2n + 1 \geq 5$, though the variance of the parent distribution is infinite (§ 4, Example 2).

Remark 2. Theorem 1 states only some sufficient conditions under which (10) and (11) are true. For a Laplace parent distribution, $f'(\xi)$ does not exist, yet (10) and (11) hold (§ 4, Example 1).

The above theorem provides a justification, at least for large samples, for using $\bar{\mu}_k$ as an approximation to $\tilde{\mu}_k$. In the next section we will proceed to show that if the parent pdf satisfies some additional conditions, then satisfactory approximations can be obtained for $\tilde{\mu}_k$ for samples of smaller sizes as well.

4. Approximations.

Lemma 2. If k is real, then the following series is convergent for every positive integer $n > k$,

$$(20) \qquad \sum_{m=1}^{\infty} \int_0^1 m^k \left| 2(F - \tfrac{1}{2}) \right|^m c_n F^n (1 - F)^n \, dF .$$

Proof. Use Lemma 1 and the fact [1, p. 33] that if $a_m \geq 0$ and $m(a_m/a_{m+1} - 1)$ approaches $r > 1$, then $\sum_{m=1}^{\infty} a_m$ is convergent; or apply the Stirling's approximation, with m large, to the gamma functions obtained by putting $c = \frac{1}{2}$ in (6).

Theorem 2. Let $F(x)$ be the cdf of a given distribution and ξ and $\tilde{x}$ be respectively the median and the sample median of a sample of size $2n + 1$. Suppose that $x(F)$, the inverse function of $F(x)$, is for $0 < F < 1$ uniquely defined and equal to a convergent series of powers of $F - \frac{1}{2}$; let

$$(21) \qquad x(F) - \xi = \sum_{m=1}^{\infty} a_m (F - \tfrac{1}{2})^m .$$

Write

$$(22) \qquad S_m = \sum_{r=1}^{m} a_r (F - \tfrac{1}{2})^r , \quad \text{and} \quad R_m = \sum_{r=m+1}^{\infty} a_r (F - \tfrac{1}{2})^r .$$

If there exists a sequence $\{b_m\}$ such that

$$(23) \qquad \sum_{m=1}^{\infty} (a_m/b_m)^2 < \infty ,$$

$$(24) \qquad \sum_{m=1}^{\infty} b_m^2 (F - \tfrac{1}{2})^{2m} < \infty, \qquad \text{for } 0 < F < 1, \quad \text{and}$$

(25) $$\sum_{m=1}^{\infty} b_m^2 \int_0^1 (F - \tfrac{1}{2})^{2m} C_n F^n(1 - F)^n \, dF < \infty ,$$

for some positive integer value n_0 of n, and $\tilde{\mu}_2$, the variance of $\tilde{x}$, is finite for $n = n_0$, then for all integers $n > n_0$,

(26) $$\tilde{\mu}_2 = \lim_{m \to \infty} \left\{ \int_0^1 S_m^2 C_n F^n(1 - F)^n \, dF - \left(\int_0^1 S_m C_n F^n(1 - F)^n \, dF \right)^2 \right\}$$

Further, if $f(x)$ is symmetric with respect to ξ, then the second term in the bracket should be omitted.

Proof. For simplicity we assume that $f(x)$ is symmetric with respect to $x = \xi$. In this case $\tilde{\mu}_1 \equiv \xi$ and

(27) $$\tilde{\mu}_2 = \int_{-\infty}^{\infty} (x - \xi)^2 C_n [F(x)]^n [1 - F(x)]^n f(x) dx = \int_{-\infty}^{a} + \int_a^b + \int_b^{\infty}$$

$$= I_1 + I_2 + I_3 ,$$

where $a < \xi$ and $b > \xi$ will be chosen later. It can be shown that

(28) $$I_1 + I_3 = O(n^{1/2} r^n) ,$$

where r is defined by (17). Choose a, b, and c such that $0 < \frac{1}{2} - F(a) = F(b) - \frac{1}{2} = c < \frac{1}{2}$. Using (21) and (22), we have

$$(29)\qquad \left| I_2 - \int_0^1 S_m^2 C_n F^n(1-F)^n dF \right| \le J_1 + J_2 + J_3 + J_4 ,$$

where

$$(30)\qquad J_1 = \int_{\frac{1}{2}+c}^{1} S_m^2 C_n F^n (1-F)^n \, dF ,$$

$$(31)\qquad J_2 = \int_{0}^{\frac{1}{2}-c} S_m^2 C_n F^n (1-F)^n \, dF ,$$

$$(32)\qquad J_3 = \int_{\frac{1}{2}-c}^{\frac{1}{2}+c} R_m^2 C_n F^n (1-F)^n \, dF ,$$

$$(33)\qquad J_4 = \int_{\frac{1}{2}-c}^{\frac{1}{2}+c} 2 \, |S_m R_m| \, C_n F^n (1-F)^n \, dF .$$

By Schwarz's inequality, we get

$$(34)\qquad J_1 + J_2 \le 6\left(\tfrac{1}{2} - c\right) \sum_{m=1}^{\infty} \left(\frac{a_m}{b_m}\right)^2 \cdot \sum_{m=1}^{\infty} b_m^2 \int_0^1 \left(F - \tfrac{1}{2}\right)^{2m} C_{n-1} F^{n-1}(1-F)^{n-1} \, dF .$$

By (23) and (25), the two series on the RHS are convergent for $n > n_0$. Hence if $n > n_0$, $J_1 + J_2$ tends to 0 as c tends to $\frac{1}{2}$.

Further, from (32),

$$(35) \qquad J_3 \leq \sum_{r=m+1}^{\infty} \left(\frac{a_r}{b_r}\right)^2 \cdot \sum_{r=m+1}^{\infty} \int_0^1 b_r^2 (F - \tfrac{1}{2})^{2r} C_n F^n (1 - F)^n \, dF$$

$$(36) \qquad J_4 \leq 2\left[\sum_{r=1}^{\infty} \left(\frac{a_r}{b_r}\right)^2 \cdot \sum_{r=m+1}^{\infty} \left(\frac{a_r}{b_r}\right)^2 \right]^{\frac{1}{2}} \cdot \sum_{r=1}^{\infty} b_r^2 \int_0^1 (F - \tfrac{1}{2})^{2r} C_n F^n (1 - F)^n \, dF \, .$$

As m tends to infinity, $J_3 + J_4$ tends to 0. Consequently, for any fixed $n > n_0$,

$$(37) \qquad \tilde{\mu}_2 = \lim_{m \to \infty} \int_0^1 S_m^2 C_n F^n (1 - F)^n \, dF \, .$$

An immediate consequence of Lemma 2 and Theorem 2 is

Theorem 3. If, in the preceding theorem, $a_m = O(2^m m^k)$ for some integer k, then (26) holds for every $n > 2k + 3$.

Proof. Choose $b_m = 2^m m^{k+1}$.

Thus we have found an approximation for

$$(38) \qquad \tilde{\mu}_2 \sim \int_0^1 S_m^2 C_n F^n (1 - F)^n \, dF \, ,$$

for all integers n which are not too small. The integral on the RHS of (38) can be evaluated by formulas given in Lemma 1. An upper bound for the error committed in such approximation is given by the sum of the RHS of (35) and (36). Finally we note that the same method can be used to obtain the moments of $\tilde{x}$ in general.

Example 1. Laplace Distribution. Let $f(x) = \frac{1}{2} e^{-|x|}$, then $F(x) = 1 - \frac{1}{2} e^{-x}$ if $x \geq 0$, and $F(x) = \frac{1}{2} e^{x}$ if $x \leq 0$. Hence

$$(39) \qquad \tilde{\mu}_2 = 2 \int_{\frac{1}{2}}^{1} x^2 C_n F^n (1 - F)^n \, dF \, .$$

If $\frac{1}{2} \leq F < 1$, then $x = -\log 2(1 - F) = \sum_{m=1}^{\infty} m^{-1} [2(F - \frac{1}{2})]^m$. So $a_m = 2^m m^{-1}$. It follows that for $n > 1$,

$$(40) \qquad \tilde{\mu}_2 = \lim_{m \to \infty} 2 \int_{\frac{1}{2}}^{1} \Big(\sum_{r=1}^{m} r^{-1} [2(F - \tfrac{1}{2})]^r \Big)^2 C_n F^n (1 - F)^n \, dF$$

$$(41) \qquad = \sum_{m=1}^{\infty} w_m \int_0^1 \left| 2(F - \tfrac{1}{2}) \right|^{m+1} C_n F^n (1 - F)^n \, dF \, ,$$

where [1, p. 84],

$$(42) \qquad w_m = \sum_{r=1}^{m} r^{-1} (m - r + 1)^{-1} = 2(m + 1)^{-1} \sum_{r=1}^{m} r^{-1} \, .$$

If we use

$$(43) \qquad \tilde{\mu}_2 \sim \sum_{m=1}^{2k-1} w_m \int_0^1 \left| 2(F - \tfrac{1}{2}) \right|^{m+1} C_n F^n (1 - F)^n \, dF \, ,$$

then the error committed is bounded by

$$(43a) \qquad 2\pi^{-1/2} w_{2k} (1 + \tfrac{1}{2n})(n + k)^{-1/2} \, \frac{2 \cdot 4 \cdots 2k}{(2n+1)(2n+3)\ldots(2n+2k-1)} \, .$$

In deriving (43a), we used the facts that w_m is a monotonically decreasing sequence

of m [1, p. 85] and the Wallis product [7, p. 385]

$$\frac{1 \cdot 3 \cdots (2n-1)}{2 \cdot 4 \cdots (2n)}\, n^{1/2} \longrightarrow \pi^{-1/2}$$

is a monotonically increasing sequence of n. Similarly if

$$(44) \qquad \tilde{\mu}_2 \sim \sum_{m=1}^{2k} w_n \int_0^1 \left| 2(F - \tfrac{1}{2}) \right|^{m+1} C_n F^n (1 - F)^n \, dF \, ,$$

then the error is bounded by

$$(44a) \qquad 2w_{2k+1}(1 + \frac{1}{2n}) \frac{1 \cdot 3 \cdots (2k+1)}{(2n+1)(2n+3)\ldots(2n+2k+1)} \, .$$

Example 2. Cauchy distribution. Let $f(x) = 1/\pi(1 + x^2)$, then $F(x) = \pi^{-1}[\tan^{-1}x + \frac{\pi}{2}]$, for $-\infty < x < \infty$, so $x(F) = \tan \pi(F - \frac{1}{2})$ for $0 < F < 1$. It can be shown that the variance of the sample median of a sample of size $2n + 1 \geq 5$ is finite:

$$(45) \qquad \tilde{\mu}_2 = \int_0^1 \tan^2 \pi(F - \tfrac{1}{2})\, C_n F^n (1 - F)^n \, dF \, .$$

It is known [7, pp. 204, 237] that

$$(46) \qquad \tan x = \sum_{m=1}^{\infty} (-1)^{m-1} \frac{2^{2m}(2^{2m} - 1)}{(2m)!} B_{2m} x^{2m-1} \, , \qquad \text{for } |x| < \frac{\pi}{2} \, ,$$

where

$$B_{2m} = 2(-1)^{m-1} \frac{(2m)!}{(2\pi)^{2m}} \sum_{r=1}^{\infty} r^{-2m} .$$

We see that $a_m = O(2^m)$, hence by Theorem 3, (37) holds if $n > 3$.

5. Normal distribution. Throughout this section, $f(x) = (2\pi)^{-1/2} e^{-x^2/2}$ and

$F(x) = \int_{-\infty}^{x} f(t)\,dt$, and $x(F)$ is the inverse function of $F(x)$. No simple general form of the derivatives of $x(F)$ at $F = \frac{1}{2}$ is known. But the first few derivatives of $x(F)$ can be obtained by direct differentiation, e.g.,

$$\frac{dx}{dF} = \frac{1}{f(x)}, \quad \frac{d^2x}{dF^2} = \frac{1}{f(x)} \frac{d}{dx}\left(\frac{1}{f(x)}\right), \ldots .$$

For finite m and $0 < F < 1$, let

$$(47) \qquad x(F) = a_1(F - \tfrac{1}{2}) + a_2(F - \tfrac{1}{2})^2 + \ldots + a_m(F - \tfrac{1}{2})^m + R_{m+1}(F - \tfrac{1}{2})^{m+1} ,$$

then

$$(48) \qquad a_2 = a_4 = \ldots = 0 ,$$

$$a_1 = (2\pi)^{1/2}, \quad a_3 = (2\pi)^{3/2}/3! , \quad a_5 = 7(2\pi)^{5/2}/5! , \ldots ,$$

$$R_5 = \frac{(2\pi)^{5/2}}{5!} \left[(7 + 46x^2 + 24x^4)e^{5x^2/2}\right]_{F_\theta}, \ldots ,$$

where $\left[g(x)\right]_{F_\theta} = g\left[x(F_\theta)\right]$, $F_\theta = \frac{1}{2} + \theta(F - \frac{1}{2})$ and $0 \le \theta \le 1$.

A. A Lower Bound. Take the integral (39), let the range of integration be divided into two: $\frac{1}{2}$ to $\frac{1}{2} + c$, and $\frac{1}{2} + c$ to 1, where $0 < c < \frac{1}{2}$. If we neglect the last integral; in the first integral, replace $x(F)$ by its expansion (47) with $m = 6$, and then neglect all terms containing the remainder term R_7 (which is non-negative), and finally let c approach $\frac{1}{2}$ and use Lemma 1, we are able to obtain a lower bound for the variance $\tilde{\mu}_2$ of $\tilde{x}$ of a normal parent distribution with unit variance, i.e.,

(49) $$\tilde{\mu}_2 \geq \lambda_2 \left[1 + \frac{\pi}{2(2n+5)} + \frac{13\pi^2}{24(2n+5)(2n+7)} \right],$$

where

(50) $$\lambda_2 = \pi/2(2n + 3).$$

Incidentally, for $n \geq 3$, we may expand the RHS of (49) into a power series in $(2n + 1)^{-1}$ and obtain an approximation for

(51) $$\tilde{\mu}_2 \sim \bar{\mu}_2 \left[1 - (2 - \tfrac{\pi}{2})(2n + 1)^{-1} - (3\pi - 4 - 13\pi^2/24)(2n + 1)^{-2} + \dots \right],$$

where $\bar{\mu}_2$ is given by (2). In terms of standard deviations, and with the numerical values of the coefficients computed, (51) is equivalent to

(52) $$\tilde{\mu}_2^{1/2} \sim \bar{\mu}_2^{1/2} \left[1 - (.2146)(2n + 1)^{-1} - (.0806)(2n + 1)^{-2} + \dots \right].$$

This agrees with a formula obtained by K. Pearson for the same purpose [8, p.363].

B. Approximations for large samples.

$$(53)\qquad \tilde{\mu}_2 = 2\int_0^{\infty} x^2 c_n [F(x)]^n [1 - F(x)]^n f(x)dx = 2\int_0^{a} + 2\int_a^{\infty} = I_1 + I_2, \text{ say.}$$

Since $F(x)[1 - F(x)] \leq F(a)[1 - F(a)]$ if $x \geq a \geq 0$ and $c_n \leq (2\pi)^{-1/2}[1 + (2n)^{-1}](2n + 1)^{1/2}2^{2n+1}$, it follows that

$$(54)\qquad I_2/\lambda_2 \leq (2/\pi)^{3/2}(1 + 3/2n)(2n + 1)^{3/2}[4F(a)\,(1-F(a))]^n 2\int_a^{\infty} x^2(2\pi)^{-\frac{1}{2}}e^{-\frac{x^2}{2}}dx.$$

In I_1, use F as the independent variable, and replace $x(F)$ by $a_1(F - \frac{1}{2}) + a_3(F - \frac{1}{2})^3 + M_5(F - \frac{1}{2})^5$ where $M_5 = \underset{0\leq x\leq a}{\text{Max}}\ R_5 = (2\pi)^{5/2}A/5!$ and $A = (7+46a^2+24a^4)e^{5a^2/2}$, then

$$(55)\qquad I_1/\lambda_2 \leq 1 + \frac{\pi}{2(2n+5)} + \left(\frac{\pi}{2}\right)^2\left(\frac{2A}{5!} + \frac{1}{3!^2}\right)\frac{3\cdot 5}{(2n+5)(2n+7)}$$
$$+ \left(\frac{\pi}{2}\right)^3 \frac{2A}{3!5!}\ \frac{3\cdot 5\cdot 7}{(2n+5)(2n+7)(2n+9)}$$
$$+ \left(\frac{\pi}{2}\right)^4 \left(\frac{A}{5!}\right)^2 \frac{3\cdot 5\cdot 7\cdot 9}{(2n+5)\ \cdots\ (2n+11)}\ .$$

Combining (49), (53), (54), and (55), we conclude:

(56) $\tilde{\mu}_2 \sim$ First Approximation: $\lambda_2 = \dfrac{\pi}{2(2n + 3)}$

Second Approximation: $\lambda_2\left[1 + \dfrac{\pi}{2(2n + 5)}\right]$.

If the second approximation is used, an upper bound for the proportional error (defined to be $|$ (True value / Approximation)$|- 1$) is given by the sum of the RHS

of (54) and the last three terms of that of (55). If the first approximation is used, then there is an additional error $\pi/2(2n+5)$, the second term in the bracket of the second approximation.

The following table is given for illustration. We choose successively for a: .35, .50, .65, .75. It is to be noted that the RHS of (54) is a decreasing function of n for, e.g., $n \geq 25$, $a = .75$. The RHS of (55) is obviously also a decreasing function of n. Therefore what Table 1 means is that: e.g., for sample sizes ≥ 51 (not just $= 51$), if the first approximation is used, then the proportional error is $\leq .092$, or explicitly: $1 \leq \mu_2/\lambda_2 \leq 1.092$.

TABLE 1

Proportional Error

Sample Size	First Approximation	Second Approximation
501	3.2×10^{-3}	6.8×10^{-5}
201	8.5×10^{-3}	7.8×10^{-4}
101	2.2×10^{-2}	6.9×10^{-3}
51	9.2×10^{-2}	6.3×10^{-2}

6. Normal distribution -- a different approach. In this section a different method is used to derive upper and lower bounds for the variance of $\tilde{x}$ of a normal parent distribution with unit variance. We state

Theorem 4. Let $\tilde{x}$ and $\tilde{\mu}_2$ be respectively the sample median and its variance of a sample of size $2n + 1$ drawn from a normal distribution with unit variance. If $\bar{\mu}_2 = \pi/2(2n + 1)$ is the variance of the asymptotic distribution of $\tilde{x}$, then

$$(57) \qquad B_n\left(1 - \frac{1}{2n + 2}\right)^{3/2} \leq \tilde{\mu}_2/\bar{\mu}_2 \leq B_n\left(1 + \frac{1}{2n}\right)^{3/2}$$

where $B_n = C_n(\frac{1}{2})^{2n+1}(2\pi)^{1/2}/(2n + 1)^{1/2}$, and $C_n = (2n + 1)!/(n!\ n!)$. Further, for all practical purpose and $n \geq 4$,

$$(58) \qquad 1 + \frac{1}{8n} - \frac{7n+3}{24n^2(2n+1)} < B_n < 1 + \frac{1}{8n} + \frac{1}{16n(8n-1)} ,$$

or

$$(59) \qquad B_n \sim 1 + \frac{1}{8n} .$$

Proof. By using the following transformations consecutively,

$$(60) \qquad u = F(y) , \quad v = u - \frac{1}{2}$$

where $F(y) = \int_{-\infty}^{y} (2\pi)^{-1/2} e^{-x^2/2} dx$, we obtain

$$(61) \qquad \tilde{\mu}_2 = 2C_n(\tfrac{1}{2})^{2n} \int_0^{1/2} y^2(1 - 4v^2)^n dv .$$

Let

(62) $$v = \frac{1}{2}(1 - e^{-t^2/(2n+1)})^{1/2} ,$$

then

(63) $$\tilde{\mu}_2 = 2B_n \int_0^{\infty} (2\pi)^{-1/2}\, y^2\, e^{-(n+1)t^2/(2n+1)}\, h_1(t/(2n+1)^{1/2})\, dt ,$$

where $h_1(t) = t(1 - e^{-t^2})^{-1/2} \geq 1$ for all $t \geq 0$. Further, it is known [10] that

(64) $$\int_0^{y} (2\pi)^{-1/2}\, e^{-x^2/2}\, dx \leq \frac{1}{2}(1 - e^{-2y^2/\pi})^{1/2} .$$

Using (64), it can be shown that $y^2 \geq \bar{\mu}_2 t^2$. Therefore it follows from (63) that

(65) $$\tilde{\mu}_2 \geq B_n\left(1 - \frac{1}{2n + 2}\right)^{3/2} \bar{\mu}_2 .$$

On the other hand, we have from (63),

(66) $$\tilde{\mu}_2 = 2B_n\bar{\mu}_2 \int_0^{\infty} (2\pi)^{-1/2} t^2\, e^{-nt^2/(2n+1)}$$

$$\cdot \{(2/\pi) y^2 h_2(t/(2n+1)^{1/2}) dt\} ,$$

where $h_2(t) = e^{-t^2}/t(1 - e^{-t^2})^{1/2}$. If we can show that $(2/\pi) y^2 h_2(t/(2n+1)^{1/2}) \leq 1$ for all $t \geq 0$, then

(67) $$\tilde{\mu}_2 \leq B_n(1 + \frac{1}{2n})^{3/2} \bar{\mu}_2 .$$

Now $y^2 h_2(t/(2n + 1)^{1/2}) \leq g_0(y)$ where

(68) $$g_0(y) = y^2(1 - 4v^2)/4v^2 .$$

It can be seen that $\lim_{y \to 0} g_0(y) = \pi/2$. Hence it suffices for our purpose to show that $g_0(y)$ is decreasing. Let " ' " denote differentiation with respect to y. Then,

(69) $$g_0'(y) = (y/2v^3)\, g_1(y) ,$$

where

(70) $$g_1(y) = v(1 - 4v^2) - y\, v' .$$

(71) $$g_1'(y) = g_2(y)\, v' , \text{ where } g_2(y) = y^2 - 12v^2 ,$$

$$g_2'(y) = (12/\pi)e^{-y^2}\, g_3(y) , \text{ where}$$

$$g_3(y) = (\pi/6)y\, e^{y^2} - e^{y^2/2} \int_0^y e^{-x^2/2}\, dx .$$

It is known [10] that

(72) $$e^{y^2/2} \int_0^y e^{-x^2/2}\, dx = \sum_{n=0}^{\infty} y^{2n+1}/1 \cdot 3 \cdots (2n+1) .$$

Hence $g_3(y) = \sum_{n=0}^{\infty} \left(\frac{\pi}{6n!} - \frac{1}{1\cdot 3\cdots(2n+1)} \right) y^{2n+1}$. It can be shown, by a similar argument used in [10] for a similar purpose, that

$$(73) \qquad g_3(y) = y^3 [a_0 y^{-2} + a_1 + a_2 y^2 + \dots]$$

where $a_0 < 0$ and $a_i > 0$, $i = 1, 2, \dots$. Hence there exists a $y_0 > 0$ such that $g_3(y) \leq 0$ if $0 \leq y \leq y_0$ and $g_3(y) \geq 0$ if $y \geq y_0$. So as y increases from 0 to ∞, $g_2(y)$ decreases steadily from 0 to a minimum and then increases steadily to ∞. Consequently $g_1'(y)$ first decreases steadily and then increases steadily. As $\lim_{y \to 0} g_1(y) = \lim_{y \to \infty} g_1(y) = 0$, it becomes clear that $g_1(y) \leq 0$ for all $y \geq 0$. Therefore $g_0(y)$ is a decreasing function of y. This completes the proof.

Finally we note that (58) is obtained by using $n! \sim (2\pi)^{\frac{1}{2}} n^{n+1/2} e^{-n+(12n)^{-1}}$.

Remark 1. The lower bound for $\tilde{\mu}_2$ given by (49) is better than the one given by (57) if we use (59) for B_n. This is so even if the last term at the RHS of (49) is omitted. For

$$(74) \qquad \frac{\pi}{2(2n+3)} + \frac{\pi^2}{4(2n+3)(2n+5)} = \frac{\pi}{2(2n+1)} \left[1 - \frac{(8-2\pi)n + 20 - \pi}{2(2n+3)(2n+5)} \right].$$

Now if $n \geq 2$, the last term in the bracket of (74) is smaller in absolute value than $(2n+2)^{-1}$ and $(1 + \frac{1}{8n}) [1 - 1/(2n+2)]^{\frac{1}{2}} \leq 1$. Therefore the quantity in the bracket of (74) is greater than

$$1 - \frac{1}{2n+2} \geq (1 + \frac{1}{8n}) (1 - \frac{1}{2n+2})^{\frac{3}{2}} .$$

For $n = 1$, direct comparison shows also that the LHS of (74) is greater than that of (57).

Remark 2. Since both lower bounds for $\tilde{\mu}_2$, (49) and (57) are smaller than $\bar{\mu}_2$, while the upper bound is greater than $\bar{\mu}_2 (1 + 1/2n)$, we cannot be sure, just by comparing these bounds, that in using $\bar{\mu}_2$ as an approximation to $\tilde{\mu}_2$, the proportional error is less than $1/2n$. Therefore the second approximation given by (56) is better than $\bar{\mu}_2$, for large samples. (Table 1).

7. Laplace distribution. We shall now employ the same technique, used in § 6, to derive upper and lower bounds for the variance $\tilde{\mu}_2$ of the sample median $\tilde{x}$ of a sample of size $2n + 1$ drawn from a Laplace distribution with pdf

$$f(x) = \frac{1}{2} e^{-|x|} . \tag{75}$$

Clearly, the variance in this case of the asymptotic distribution of $\tilde{x}$ is

$$\bar{\mu}_2 = \frac{1}{2n+1} . \tag{76}$$

We state

Theorem 5. If $\tilde{\mu}_2$ and $\bar{\mu}_2$ are as defined above, then

(77) $$B_n\left(1 - \frac{1}{2n+2}\right)^{\frac{3}{2}} \leq \tilde{\mu}_2/\bar{\mu}_2 \leq 1.51 \; B_n\left(1 + \frac{1}{2n}\right)^{\frac{3}{2}} ,$$

where B_n is given by (57) and (59).

Proof. It can be seen easily that $\tilde{\mu}_2$ is equal to the RHS of (63) if v and t satisfy (62) and

(78) $$v = \int_0^y \frac{1}{2} e^{-x} \, dx.$$

We proved [4] that for all $y \geq 0$,

(79) $$v \leq \frac{1}{2}\left(1 - e^{-y^2}\right)^{\frac{1}{2}} .$$

From (62) and (79), we have $y^2 \geq \bar{\mu}_2 t^2$. Hence it follows from (63) that $\tilde{\mu}_2/\bar{\mu}_2$ has a lower bound given by (77).

Further, from (63)

(80) $$\tilde{\mu}_2 = 2B_n\bar{\mu}_2 \int_0^{\infty} (2\pi)^{-\frac{1}{2}} t^2 e^{-nt^2/(2n+1)} \{ y^2 h_2(t/(2n+1)^{\frac{1}{2}}) \, dt \},$$

where $h_2(t)$ is given by (66). We say that

(81) $$y^2 h_2(t / (2n+1)^{\frac{1}{2}}) \leq 1.51 .$$

If this is true, then the RHS of (77) is an upper bound of $\tilde{\mu}_2/\bar{\mu}_2$. Thus the proof is completed.

To establish (81), we introduce, as in (68),

(82) $$g_0(y) = y^2(1 - 4v^2)/4v^2$$

where y and v satisfy (78). For all $y \geq 0$, $g_0(y)$ is not smaller than the LHS of (81) and

(83) $$\text{Lim}\; g_0(y) = \begin{matrix} 1 \\ 0 \end{matrix} \quad \text{as} \quad y \longrightarrow \begin{matrix} 0 \\ \infty \end{matrix} .$$

Let "'" denote differentiation with respect to y, then

(84) $$g_0'(y) = \frac{y}{2v^3}\, g_1(y) \quad , \quad \text{where}$$

(85) $$g_1(y) = v - 4v^3 - \frac{1}{2}\, y\, e^{-y} \quad .$$

(86) $$g_1'(y) = \frac{1}{2}\, e^{-y}\, g_2(y) \quad , \quad \text{where}$$

(87) $$g_2(y) = -\, 12v^2 + y \quad .$$

(88) $$g_2'(y) = -\, 12\, v\, e^{-y} + 1 = g_3(y)$$

(89) $$g_3'(y) = 12\, e^{-y}\, (\frac{1}{2} - e^{-y}) \quad .$$

If f(x) is a function of x, and if as x increases from 0 to oo, f(x) varies from, e.g., positive to negative, and then back to positive, we will write, for simplicity, As x: $0 \longrightarrow \infty$, f(x): +,-,+. Now

$$g_3'(y) \gtrless 0 \text{ according as } y \gtrless \log 2, \tag{90}$$

and $g_3(0) = g_3(\infty) = 1$, and $g(\log 2) = -\frac{1}{2}$. So as y: $0 \longrightarrow \infty$, $g_3(y)$: +,-,+. Now $g_2(0) = 0$, while $g_2(\infty) = \infty$. We say that as y: $0 \longrightarrow \infty$, $g_2(y)$: +,-,+. Otherwise $g_2(y) \geq 0$ for all $y \geq 0$, so $g_1'(y) \geq 0$ and $g_1(y) \geq 0$ for all $y \geq 0$, as $g_1(0) = 0$. Hence $g_0(y)$ is steadily increasing. This, however, contradicts (83). It follows that as y: $0 \longrightarrow \infty$, $g_1'(y)$: +,-,+. Now $g_1(0) = g_1(\infty) = 0$, hence as y: $0 \longrightarrow \infty$, $g_1(y)$: +,-. Therefore we conclude that as y: $0 \longrightarrow \infty$, $g_0(y)$ increases steadily from 1 to a maximum, and then decreases steadily to 0. To find the maximum of $g_0(y)$, we first solve $g_1(y) = 0$, which is equivalent to $2v(1+2v) - y = 0$. Using table [12], we obtain an approximate solution $y = 1.15$. The maximum of $g_0(y)$ is then found to be 1.51.

Remark. The variance of the sample mean (of a sample of size 2n+1) drawn from a Laplace distribution with pdf given by (75) is 2/(2n+1). It follows, from Theorem 5, that the sample median has smaller variance than the sample mean for sample size $2n+1 \geq 7$. In a recent paper, A. E. Sarhan [11] found that for sample sizes equal to 2, 3, 4, and 5, the variance of sample median is also smaller than that of the sample mean.

References

[1] T. J. I'a. Bromich, An Introduction to the theory of Infinite Series, MacMillan and Co., London, 1908.

[2] G. W. Brown and J. W. Tukey, Some distributions of sample means", Ann. Math. Stat., Vol. 17 (1946), pp. 1-12.

[3] J. H. Cadwell, "The distribution of quantiles of small samples", Biometrika, Vol. 39 (1952), pp. 207-211.

[4] J. T. Chu, "On the distribution of the sample median", to be published.

[5] W. Feller, "On the normal approximation to the binomial distribution", Ann. Math. Stat., Vol. 16 (1945), pp. 319-329.

[6] T. Hojo, "Distirbution of the median, etc.", Biometrika, Vol. 23 (1931), pp. 315-360.

[7] K. Knopp, Theory and Application of Infinite Series, Hafner Publishing Company, New York.

[8] K. Pearson, "On the standard error of the median, etc.", Biometrika 23 (1931), pp. 361-363.

[9] K. Pearson, "On the mean character and variance of a ranked individual, etc." Biometrika, Vol. 23 (1931), pp. 364-397.

[10] G. Pólya, "Remarks on computing the probability integral in one and two dimensions", Proceedings of the Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, 1949, pp. 63-78.

[11] A. E. Sarhan, "Estimation of the mean and standard deviation by order statistics", Ann. Math. Stat., Vol. 25 (1954), pp. 317-328.

[12] Tables of the Exponential Function e^x, Applied Mathematics Series, 14, National Bureau of Standards, Washington, D. C., 1951.